Mathematical modeling of textures: application to color image decomposition with a projected gradient algorithm

Vincent Duval, Jean-François Aujol, Luminita Vese

A projected gradient algorithm for color image decomposition

Vincent Duval [1] & Jean-François Aujol [2] & Luminita Vese [3]
[1] Institut TELECOM, TELECOM ParisTech, CNRS UMR 5141
vduval@enst.fr
[2] CMLA, ENS Cachan, CNRS, UniverSud
Jean-Francois.Aujol@cmla.ens-cachan.fr
[3] UCLA, Mathematics Department
lvese@ucla.math.edu

Abstract: In this paper, we are interested in color image processing, and in particular color image decomposition. The problem of image decomposition consists in splitting an original image f into two components u and v. u should contain the geometric information of the original image, while v should be made of the oscillating patterns of f, such as textures. We propose here a scheme based on a projected gradient algorithm to compute the solution of various decomposition models for color images or vector-valued images. We provide a direct convergence proof of the scheme, and we give some analysis on color texture modeling.

Key-words: Color image decomposition, projected gradient algorithm, color texture modeling.

1 Introduction

Since the seminal work by Rudin et al [50], total variation based image restoration and decomposition has drawn a lot of attention (see [25, 6, 48, 4] and references therein for instance). f being the original image, we are interested in minimizing energies of the type:

$$\int |Du| + \mu \|f - u\|_T^k \tag{1}$$

$\int |Du|$ stands for the total variation; in the case when u is regular, then $\int |Du| = \int |\nabla u| \, dx$. $\|.\|_T$ stands for a norm which captures the noise and/or the textures of the original image f (in the sense that it is not too large for such features) and k is a positive exponent.

The most basic choice for $\|.\|_T$ is the L^2 norm, and $k = 2$. From a Bayesian point of view, this is also the norm which appears naturally when assuming that the image f has been corrupted by some Gaussian white noise (see e.g. [25]). However, since the book by Y. Meyer [44], other spaces have been considered for modeling oscillating patterns such as textures or noise. The problem of image decomposition has been a very active field of research during the last past five years. [44], was the inspiration source of many works, e.g. [54, 49, 7, 5, 53, 8, 15, 23, 33, 58, 59, 60, 38, 9, 37, 41]. Image decomposition consists in splitting an original image f into two components, u and $v = f - u$. u is supposed to contain the geometrical component of the original image (it can be seen as some kind of

sketch of the original image), while v is made of the oscillatory component (the texture component in the case when the original image f is noise free).

In this paper, we are concerned with color image processing. While some authors deal with color images using a Riemannian framework, like G. Sapiro and D. L. Ringach [51] or N. Sochen, R. Kimmel and R. Malladi [52], others combine a functional analysis viewpoint with the Chromaticity-Brightness representation [10]. The model we use is more basic: it is the same as the one used in [19] (and related with [18]). Its advantage is to have a rich functional analysis interpretation. Note that in [55], the authors also propose a cartoon + texture color decomposition and denoising model inspired from Y. Meyer, using the vectorial versions of total variation and approximations of the space $G(\Omega)$ for textures (to be defined later); unlike the work presented here, they use Euler-Lagrange equations and gradient descent scheme for the minimization, which should be slower than by projection methods.

Here, we give some insight into the definition of a texture space for color images. In [13], a TV-Hilbert model was proposed for image restoration and/ or decomposition:

$$\int |Du| + \mu \|f - u\|_{\mathcal{H}}^2 \tag{2}$$

where $\|.\|_{\mathcal{H}}$ stands for the norm of some Hilbert space $\mathcal{H}$. This is a particular case of problem (1). Due to the Hilbert structure of $\mathcal{H}$, there exist many different methods to minimize (2), such as a projection algorithm [13]. We extend (2) to the case of color images.

From a numerical point of view, (1) is not straightforward to minimize. Depending on the choice for $\|.\|_T$, the minimization of (1) can be quite challenging. Nevertheless, even in the simplest case when $\|.\|_T$ is the L^2 norm, handling the total variation term $\int |Du|$ needs to be done with care. The most classical approach consists in writing the associated Euler-Lagrange equation to problem (1). In [50], a fixed step gradient descent scheme is used to compute the solution. This method has on the one hand the advantage of being very easy to implement, and on the other hand the disadvantage of being quite slow. To improve the convergence speed, quasi-Newton methods have been proposed [22, 56, 34, 1, 26, 46, 47]. Iterative methods have proved successful [17, 32, 15]. A projected-subgradient method can be found in [28].

Duality based schemes have also drawn a lot of attention to solve (1): first by Chan and Golub in [24], later by A. Chambolle in [20] with a projection algorithm, and then generalized in [29]. This projection algorithm has recently been extended to the case of color images in [19]. Second order cone programming ideas and interior point methods have proved interesting approaches [39, 36]. Recently, it has been shown that graph cuts based algorithms could also be used [21, 31]. Finally, let us notice that it is shown in [57, 12] that Nesterov's scheme [45] provides fast algorithms for minimizing (1).

Another variant of Chambolle projection algorithm [20] is to use a projected gradient algorithm [21, 12]. Here we have decided to use this approach which has both advantages of being easy to implement and of being quite efficient.

The plan of the paper is the following. In Section 2, we define and provide some analysis about the spaces we consider in the paper. In Section 3, we extend the TV-Hilbert model originally introduced in [13] to the case of color images. In Section 4, we present a projected gradient algorithm to compute a minimizer of problem (2). This projected gradient algorithm has first been proposed by A. Chambolle in [21] for total variation regularization. A proof of convergence was given in [12] relying on optimization results by Bermudez and Moreno [16]. We derive here a simple and direct proof of convergence. In Section 5, we apply this scheme to solve various classical denoising and decomposition problems. We illustrate our approach with many numerical examples.

2 Definitions and properties of the considered color spaces

In this section, we introduce some notations, and we provide some analysis of the functional analysis spaces we consider to model color textures.

2.1 Introduction

Let Ω be a Lipschitz convex bounded open set in $\mathbb{R}^2$. We model color images as $\mathbb{R}^M$-valued functions defined on Ω. The inner product in $L^2(\Omega, \mathbb{R}^M)$ is denoted as:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle_{L^2(\Omega,\mathbb{R}^M)} = \int_\Omega \sum_{i=1}^{M} u_i v_i.$$

For a vector $\xi \in \mathbb{R}^M$, we define the norms:

$$|\xi|_1 = \sum_{i=1}^{M} |\xi_i|, \tag{3}$$

$$|\xi|_2 = \sqrt{\sum_{i=1}^{M} \xi_i^2}, \tag{4}$$

$$|\xi|_\infty = \max_{i=1\ldots M} |\xi_i|. \tag{5}$$

We will sometimes refer to the space of zero-mean functions in $L^2(\Omega, \mathbb{R}^M)$ by V_0:

$$V_0 = \{\boldsymbol{f} \in L^2(\Omega, \mathbb{R}^M), \int_\Omega \boldsymbol{f} = 0\}.$$

We say that a function $\boldsymbol{f} \in L^1(\Omega, \mathbb{R}^M)$ has bounded variation if the following quantity is finite:

$$\begin{aligned} |\boldsymbol{f}|_{TV} &= \sup \left\{ \int_\Omega \sum_{i=1}^{M} f_i \mathrm{div} \xi_i,\ \boldsymbol{\xi} \in P \right\} \\ &= \sup_{\boldsymbol{\xi} \in P} \langle \boldsymbol{f}, \mathrm{div}\ \boldsymbol{\xi} \rangle_{L^2(\Omega,\mathbb{R}^M)} \end{aligned}$$

where P is a subset of $\boldsymbol{\xi} \in C_c^1(\Omega, \mathbb{R}^{2\times M})$. This quantity is called the total variation. For more information on its properties, we refer the reader to [3]. The set of functions with bounded variation is a vector space classically denoted by $BV(\Omega, \mathbb{R}^M)$.

In this paper, we will consider the following set of test-functions, which is the classical choice ([3], [6]) :

$$\mathcal{B} = \{\xi \in C_c^1(\Omega, \mathbb{R}^{2\times M}) / \forall x \in \Omega, |\boldsymbol{\xi}(x)|_2 \leq 1\}.$$

Then, for $\boldsymbol{f}$ smooth enough, the total variation of $\boldsymbol{f}$ is: :

$$|\boldsymbol{f}|_{TV} = \int_\Omega \sqrt{\sum_{i=1}^{M} |\nabla f_i|^2} dx.$$

As X. Bresson and T. Chan notice in [19], the choice of the set $\mathcal{B}$ is crucial. If one chooses:

$$\mathcal{B} = \{\xi \in C_c^1(\Omega, \mathbb{R}^{2\times M}) / \forall x \in \Omega, |\boldsymbol{\xi}(x)|_\infty \leq 1\}$$

then one has :

$$|\boldsymbol{f}|_{TV} = \sum_{i=1}^{M} \int_\Omega |\nabla f_i| dx = \sum_{i=1}^{M} |f_i|_{TV}.$$

These two choices are mathematically equivalent and define the same BV space, but in practice the latter induces no coupling between the channels, which gives visual artifacts in image processing.

2.2 The color $G(\Omega)$ space

The $G(\mathbb{R}^2)$ space was introduced by Y. Meyer in [44] to model textures in grayscale images. It was defined as div $\left(L^\infty(\mathbb{R}^2)\right)$, but one could show that this space was equal to $W^{-1,\infty}(\mathbb{R}^2)$ (the dual of $W^{1,1}(\mathbb{R}^2)$). For the generalization to color images, we will adopt the framework of ([11], [5]; the color space $G(\Omega)$ is also used in [55], as a generalization of [54] to color image decomposition and color image denoising).

Definition 2.1. The space $G(\Omega)$ is defined by :

$$G(\Omega) = \{\boldsymbol{v} \in L^2(\Omega, \mathbb{R}^M) / \exists \vec{\boldsymbol{\xi}} \in L^\infty(\Omega, (\mathbb{R}^2)^M), \forall i = 1, \ldots, M,\ v_i = \text{div } \vec{\xi}_i \text{ and } \vec{\xi}_i \cdot \vec{N} = 0 \text{ on } \partial\Omega\}$$

(where $\vec{\xi}_i \cdot \vec{N}$ refers to the normal trace of $\vec{\xi}_i$ over $\partial\Omega$). One can endow it with the norm :

$$\|\boldsymbol{v}\|_G = \inf\{\|\vec{\boldsymbol{\xi}}\|_\infty,\ \forall i = 1, \ldots, M,\ v_i = \text{ div } \xi_i, \vec{\xi}_i \cdot \vec{N} = 0 \text{ on } \partial\Omega\}$$

with $\|\vec{\boldsymbol{\xi}}\|_\infty = \sup \text{ ess } \sqrt{\sum_{i=1}^{M} |\vec{\xi}_i|^2}$.

The following result was proved in [11] for grayscale images: it characterizes $G(\Omega)$.

Proposition 2.1.

$$G(\Omega) = \left\{\boldsymbol{v} \in L^2(\Omega, \mathbb{R}^M) / \int_\Omega \boldsymbol{v} = 0\right\}.$$

Proof: Let us introduce the grayscale $G^1(\Omega)$ space:

$$G^1(\Omega) = \{v \in L^2(\Omega) / \exists \vec{\xi} \in L^\infty(\Omega, \mathbb{R}^2), v = \text{div } \vec{\xi} \text{ and } \vec{\xi} \cdot \vec{N} = 0 \text{ on } \partial\Omega\}.$$

By the result proved in [11], Proposition 4.1.1, we have that

$$G^1(\Omega) = \left\{v \in L^2(\Omega) / \int_\Omega v = 0\right\}.$$

We now proceed component by component.

Let $\boldsymbol{v} \in L^2(\Omega, \mathbb{R}^M)$. If $\int_\Omega \boldsymbol{v} = 0$, then for each $i \in 1, \ldots, M$, $\int_\Omega v_i = 0$. The 1D result tells us that $v_i \in G^1(\Omega)$, thus there exists $\vec{\xi_i} \in L^\infty(\Omega, \mathbb{R}^2)$ with $\vec{\xi_i} \cdot \vec{N} = 0$ on $\partial\Omega$ such that $v_i = \text{ div } \vec{\xi_i}$. Then the vector $\vec{\boldsymbol{\xi}} = \begin{pmatrix} \vec{\xi_1} \\ \vdots \\ \vec{\xi_M} \end{pmatrix} \in L^\infty(\Omega, (\mathbb{R}^2)^M)$ and $\boldsymbol{v} \in G(\Omega)$.

Conversely, if $\boldsymbol{v} \in G(\Omega)$, then there exists $\vec{\boldsymbol{\xi}} \in L^\infty(\Omega, (\mathbb{R}^2)^M)$, such that for each $i \in 1, \ldots, M$, with $\vec{\xi_i} \cdot \vec{N} = 0$ on $\partial\Omega$ and that $v_i = \text{ div } \vec{\xi_i}$. By integration component by component, $\int_\Omega \boldsymbol{v} = 0$.

■

In L. Lieu's PhD thesis [42], one can find the following property (generalized here to the vectorial case):

Lemma 2.1. *For $\boldsymbol{f} \in L^2(\Omega, \mathbb{R}^M)$, let us consider the semi-norm :*

$$\|\boldsymbol{f}\|_* = \sup_{\boldsymbol{u} \in BV(\Omega, \mathbb{R}^M), |\boldsymbol{u}|_{TV} \neq 0} \frac{\langle \boldsymbol{f}, \boldsymbol{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} = \sup_{\boldsymbol{u} \in BV(\Omega, \mathbb{R}^M), |\boldsymbol{u}|_{TV} \neq 0} \frac{\sum_{i=1}^M \int_\Omega f_i u_i}{|\boldsymbol{u}|_{TV}}$$

If $\|\boldsymbol{f}\|_* < +\infty$, then $\int_\Omega \boldsymbol{f} = 0$.

Proof: Let $i \in \{1, \ldots, M\}$, and $\boldsymbol{h} \in L^2(\Omega, \mathbb{R}^M)$, $h_j = 0$ for $j \neq i$ and $h_i = 1$. Since the total variation does not change by the addition of a constant, we have:

$$\forall \boldsymbol{u} \in BV(\Omega, \mathbb{R}^M) \qquad \text{such that } |\boldsymbol{u}|_{TV} \neq 0, \forall c \in \mathbb{R},$$
$$\frac{\langle \boldsymbol{f}, \boldsymbol{u} + c\boldsymbol{h} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u} + c\boldsymbol{h}|_{TV}} = \frac{\langle \boldsymbol{f}, \boldsymbol{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} + c\frac{\langle \boldsymbol{f}, \boldsymbol{h} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} \leq \|\boldsymbol{f}\|_*$$

As c goes to $\pm\infty$, we notice that necessarily $\langle \boldsymbol{f}, \boldsymbol{h} \rangle_{L^2(\Omega)} = \int_\Omega f_i = 0$.

■

Comparing this property to Proposition 2.1, we can deduce that any function $\boldsymbol{f}$ such that $\|\boldsymbol{f}\|_* < \infty$ belongs to $G(\Omega)$. The converse is also true:

Lemma 2.2. *Let $f \in G(\Omega)$. Then $\|f\|_* < \infty$.*

Proof: Let $\boldsymbol{f} \in G(\Omega)$. Since $\int_\Omega \boldsymbol{f} = 0$, the quantity $\frac{\langle \boldsymbol{f}, \boldsymbol{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}}$ does not change with the addition of constants to $\boldsymbol{u}$. Thus

$$\begin{aligned} \sup_{\boldsymbol{u} \in BV, |\boldsymbol{u}|_{TV} \neq 0} \frac{\langle \boldsymbol{f}, \boldsymbol{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} &= \sup_{\boldsymbol{u} \in BV, \int \boldsymbol{u} = 0} \frac{\langle \boldsymbol{f}, \boldsymbol{u} \rangle_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} \\ &\leq \sup_{\boldsymbol{u} \in BV, \int \boldsymbol{u} = 0} \frac{\|\boldsymbol{f}\|_{L^2(\Omega, \mathbb{R}^M)} \|\boldsymbol{u}\|_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} \\ &\leq C\|\boldsymbol{f}\|_{L^2(\Omega)} < \infty \end{aligned}$$

where we used the Poincaré inequality : $\|\boldsymbol{u} - \boldsymbol{u}_\Omega\|_{L^2(\Omega,\mathbb{R}^M)} \leq C|u|_{TV}$.

■

The following theorem completes the above two lemmas:

Theorem 2.2. *The following equality holds :*

$$G(\Omega) = \{\boldsymbol{f} \in L^2(\Omega) / \sup_{\boldsymbol{u}\in BV(\Omega,\mathbb{R}^M),|\boldsymbol{u}|_{BV}\neq 0} \frac{\langle \boldsymbol{f}, \boldsymbol{u}\rangle_{L^2(\Omega,\mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} < +\infty\}$$

and for all function $\boldsymbol{f} \in L^2(\Omega,\mathbb{R}^M)$, $\|\boldsymbol{f}\|_* = \|\boldsymbol{f}\|_G$.

Moreover, the infimum in the definition of $\|\cdot\|_G$ *is reached.*

Proof:

(i) Let $\boldsymbol{f}$ be a function in the set on the right hand-side. Thanks to Lemma 2.1, we know that $\forall i \in \{1,\dots,M\}, \int_\Omega f_i = 0$. By Proposition 2.1, $\boldsymbol{f} \in G(\Omega)$.

Now let $\boldsymbol{u} \in BV(\Omega,\mathbb{R}^M)$ such that $|\boldsymbol{u}|_{TV} \neq 0$. By convolution with an approximation of identity, one can find a sequence $\boldsymbol{u}_n \in C^\infty(\Omega,\mathbb{R}^M) \cap W^{1,1}(\Omega;\mathbb{R}^M)$ such that $\|\boldsymbol{u} - \boldsymbol{u}_n\|_2 \to 0$ and $|\boldsymbol{u}_n|_{TV} \to |\boldsymbol{u}|_{TV}$.

Then, for all $\vec{\boldsymbol{g}}$ such that $\boldsymbol{f} = \text{div } \vec{\boldsymbol{g}}$ and $\vec{\boldsymbol{g}} \cdot \vec{N} = 0$ on $\partial\Omega$:

$$\begin{aligned}
\langle \boldsymbol{f}, \boldsymbol{u}_n\rangle_{L^2(\Omega,\mathbb{R}^M)} &= \sum_{i=1}^M \int_\Omega \text{div } \vec{\boldsymbol{g}}\ \boldsymbol{u}_n \\
&= -\int_\Omega (\sum_{i=1}^M \vec{g}_i \cdot \nabla u_{i,n}) \\
&\leq \int_\Omega |\vec{\boldsymbol{g}}||\nabla \boldsymbol{u}_n| \\
&\leq \|\vec{\boldsymbol{g}}\|_\infty |\boldsymbol{u}_n|_{TV}
\end{aligned}$$

Since $\boldsymbol{f} \in L^2(\Omega)$, we can pass to the limit in both sides of the inequality, and by construction of $\boldsymbol{u}_n$, we get:

$$\|\boldsymbol{f}\|_* \leq \|\boldsymbol{f}\|_G.$$

(ii) For the converse inequality, the proof is standard (see e.g. [40], [2]).

Let $\boldsymbol{f} \in L^2(\Omega,\mathbb{R}^M)$ such that $\sup_{\boldsymbol{u}\in BV(\Omega,\mathbb{R}^M),|\boldsymbol{u}|_{TV}\neq 0} \frac{\langle \boldsymbol{f},\boldsymbol{u}\rangle_{L^2(\Omega,\mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}} < +\infty$. Let us define :

$$T : \left\{ \begin{array}{ccc} \mathcal{D}(\bar{\Omega},\mathbb{R}^M) & \to & L^1(\Omega,\mathbb{R}^{2M}) \\ \varphi & \mapsto & \left(\frac{\partial\varphi^1}{\partial x_1}, \frac{\partial\varphi^1}{\partial x_2}, \dots, \frac{\partial\varphi^M}{\partial x_1}, \frac{\partial\varphi^M}{\partial x_2}\right) \end{array} \right.$$

To each vector $\left(\frac{\partial\varphi^1}{\partial x_1}, \frac{\partial\varphi^1}{\partial x_2}, \dots, \frac{\partial\varphi^M}{\partial x_1}, \frac{\partial\varphi^M}{\partial x_2}\right) \in T(\mathcal{D}(\bar{\Omega},\mathbb{R}^M))$, we can associate $\int_\Omega \sum_{i=1}^M f^i\varphi^i dx$ (without ambiguity since f_i has zero-mean, and if two functions have the same gradient over Ω

they only differ by a constant on the convex domain Ω). We have $\int_\Omega \sum_{i=1}^M f^i\varphi^i dx \le \|\boldsymbol{f}\|_*|\varphi|_{BV} = \|\boldsymbol{f}\|_*\|(\frac{\partial\varphi_1}{\partial x_1},\dots,\frac{\partial\varphi_M}{\partial x_2})\|_1$ thus we have defined a bounded linear form on $T(\mathcal{D}(\bar\Omega,\mathbb{R}^M))$. Using Hahn-Banach's theorem, we can extend it to $L^1(\bar\Omega,\mathbb{R}^{2M})$ with the same norm $\|\boldsymbol{f}\|_*$. Since $L^\infty(\Omega,\mathbb{R}^{2M})$ is identifiable with the dual of $L^1(\Omega,\mathbb{R}^{2M})$, there exists $g\in L^\infty(\Omega,\mathbb{R}^{2M})$ with $\|g\|_{L^\infty(\Omega,\mathbb{R}^M)} = \|\boldsymbol{f}\|_*$, such that:

$$\forall\boldsymbol{\varphi}\in\mathcal{D}(\bar\Omega,\mathbb{R}^M), \int_\Omega \sum_{i=1}^M f_i\varphi_i = -\int_\Omega \sum_{i=1}^M\sum_{j=1}^2 \frac{\partial\varphi_i}{\partial x_j} g_{i,j} = -\sum_{i=1}^M\int_\Omega \vec{g_i}\cdot\nabla\varphi_i. \tag{6}$$

This is true in particular for $\boldsymbol{\varphi}\in\mathcal{D}(\Omega,\mathbb{R}^M)$, thus $\boldsymbol{f} = \text{div}\,\vec{\boldsymbol{g}}$ in the distribution sense, and since the functions are in $L^2(\Omega,\mathbb{R}^M)$ there is equality in $L^2(\Omega,\mathbb{R}^M)$. Since div $\vec{\boldsymbol{g}}\in L^2(\Omega,\mathbb{R}^M)$, we can then consider the normal trace of $\vec{\boldsymbol{g}}$.

If $\boldsymbol{\varphi}\in\mathcal{D}(\bar\Omega,\mathbb{R}^M)$, we have by (6):

$$\sum_{i=1}^M\int_\Omega f_i\varphi_i = -\sum_{i=1}^M\int_\Omega g_i\cdot\nabla\varphi_i.$$

But on the other hand, by integration by parts :

$$\sum_{i=1}^M\int_\Omega \text{div}\ \vec{g_i}\varphi_i = -\sum_{i=1}^M\int_\Omega \vec{g_i}\cdot\nabla\varphi_i + \sum_{i=1}^M\int_{\partial\Omega}\varphi_i\vec{g_i}\cdot\vec{N}$$

The equality $\boldsymbol{f} = \text{div}\,\vec{\boldsymbol{g}}$ in $L^2(\Omega,\mathbb{R}^M)$ shows that the edge contribution vanishes for $\boldsymbol{\varphi}\in\mathcal{D}(\bar\Omega,\mathbb{R}^M)$. Thus $\vec{g_i}\cdot\vec{N} = 0$ over $\partial\Omega$.

Incidentally, we notice that the infimum in the G-norm is reached.

■

Remark 2.3. From the proof of Lemma 2.2, we can deduce that the topology induced by the G-norm on $G(\Omega)$ is coarser than the one induced by the L^2 norm. More generally, there exists a constant $C>0$ (depending only on Ω), such that:

$$\forall\boldsymbol{f}\in G(\Omega), \|\boldsymbol{f}\|_* \le C\|\boldsymbol{f}\|_{L^2(\Omega,\mathbb{R}^M)}. \tag{7}$$

In fact the G norm is strictly coarser than the L^2 norm:

Let us consider, for $m\in\mathbb{N}^*$, the sequence $\forall k = 1\dots M,\ f_m^{(k)}(x,y) = \cos mx + \cos my$ defined on $(-\pi,\pi)^M$. The vector field $\boldsymbol{\xi}^{(k)} = (\frac{1}{m}\sin(mx), \frac{1}{m}\sin(my))$ satisfies the boundary condition, and its divergence is equal to $\boldsymbol{f}_m$. As a consequence $\|\boldsymbol{f}_m\|_* \le \frac{\sqrt{2M}}{m}$ and $\lim_{m\to+\infty}\|\boldsymbol{f}_m\|_* = 0$.

Yet,

$$\begin{aligned}\|\boldsymbol{f}_m\|^2_{L^2(\Omega,\mathbb{R}^M)} &= M\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(\cos mx+\cos my)^2dxdy\\ &= 4M\pi^2.\end{aligned}$$

The sequence $\boldsymbol{f}_m$ converges to 0 for the topology induced by the G-norm, but not for the one induced by the L^2 norm .

The following result was proposed by Y. Meyer in [44]. It is still true for color images.

Proposition 2.2. *Let $\boldsymbol{f}_n, n \geq 1$ be a sequence of functions of $L^q(\Omega, \mathbb{R}^M) \cap G(\Omega)$ with the following properties :*

1. *There exists $q > 2$ and $C > 0$ such that $\|\boldsymbol{f}_n\|_{L^q(\Omega,\mathbb{R}^M)} \leq C$.*
2. *The sequence f_n converges to 0 in the distributional sense.*

Then $\|\boldsymbol{f}_n\|_G$ converges to 0 when n goes to infinity.

It means that oscillating patterns with zero mean have a small G norm.

3 Color TV-Hilbert model : presentation and mathematical analysis

3.1 Presentation

The TV-Hilbert framework was introduced for grayscale images by J.-F. Aujol and G. Gilboa in [13] as a way to approximate the BV-G model. They prove that one can extend Chambolle's algorithm to this model.

In this section we show that this is still true for color images. We are interested in solving the following problem:

$$\inf_{\boldsymbol{u}} |\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{f} - \boldsymbol{u}\|^2_{\mathcal{H}} \tag{8}$$

where $\mathcal{H} = V_0$ (the space of zero-mean functions of $L^2(\Omega, \mathbb{R}^M)$) is a Hilbert space endowed with the following norm :

$$\|\boldsymbol{v}\|^2_{\mathcal{H}} = \langle \boldsymbol{v}, K\boldsymbol{v}\rangle_{L^2(\Omega,\mathbb{R}^M)}$$

with $K : \mathcal{H} \to L^2(\Omega, \mathbb{R}^M)$

- is a bounded linear operator (for the topology induced by the $L^2(\Omega, \mathbb{R}^M)$ norm on $\mathcal{H}$)
- is symmetric positive definite

and K^{-1} is bounded on $Im(K)$.

Examples:

- The Rudin-Osher Fatemi model
  It was proposed in [50] by L. Rudin, S. Osher, and E. Fatemi for grayscale images. It was then extended to color images using different methods, for instance by G. Sapiro and D.L. Ringach [51], or Blomgren and T. Chan [18] . In [19], X. Bresson and T. Chan use another kind of color total variation, which is the one we use in this paper.

  The idea is to minimize the functional:

$$|\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{f} - \boldsymbol{u}\|^2_{L^2(\Omega,\mathbb{R}^M)}. \tag{9}$$

It is clear that the problem commutes with the addition of constants. If the (unique) solution associated to $\boldsymbol{f}$ is $\boldsymbol{u}$, then the solution associated to $\boldsymbol{f}+C$ is $(\boldsymbol{u}+C)$. As a consequence we can always assume that $\boldsymbol{f}$ has zero mean.

Then this model becomes a particular case of the TV-Hilbert model with $K = Id$.

- The OSV model
  In [49], S. Osher, A. Solé and L. Vese propose to model textures by the H^{-1} space. In order to generalize this model, we must be cautious about the meaning of our notations but it is natural to introduce the following functional :

$$\inf_{\boldsymbol{u}} |\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\int_\Omega |\boldsymbol{\nabla}\boldsymbol{\Delta}^{-1}(\boldsymbol{f}-\boldsymbol{u})|^2 \tag{10}$$

where $\boldsymbol{\Delta}^{-1}\boldsymbol{v} = \begin{pmatrix} \Delta^{-1}v_1 \\ \vdots \\ \Delta^{-1}v_M \end{pmatrix}$, $\boldsymbol{\nabla\rho} = \begin{pmatrix} \nabla\rho_1 \\ \vdots \\ \nabla\rho_M \end{pmatrix}$, $|\boldsymbol{\nabla\rho}|^2 = |\nabla\rho_1|^2 + |\nabla\rho_2|^2 + \ldots + |\nabla\rho_M|^2$ and

$$\begin{aligned}\int_\Omega |\nabla\Delta^{-1}(\boldsymbol{f}-\boldsymbol{u})|^2 = \int_\Omega \sum_{i=1}^M |\nabla\Delta^{-1}(f^i-u^i)|^2 &= -\int_\Omega \sum_{i=1}^M (f^i-u^i)\Delta^{-1}(f^i-u^i) \\ &= \langle \boldsymbol{f}-\boldsymbol{u}, -\boldsymbol{\Delta}^{-1}(\boldsymbol{f}-\boldsymbol{u})\rangle_{L^2(\Omega,\mathbb{R}^M)}.\end{aligned}$$

The inversion of the Laplacian is defined component by component. For one component, it is defined by the following proposition:

Proposition 3.1. *Let $X_0 = \{P \in H^1(\Omega,\mathbb{R}) : \int_\Omega P = 0\}$. If $v \in L^2(\Omega)$, with $\int_\Omega v = 0$, then the problem:*

$$-\Delta P = v,\ \frac{\partial P}{\partial n}|_{\partial\Omega} = 0$$

admits a unique solution in X_0.

Proof: This is a consequence of the Lax-Milgram theorem in the Hilbert space X_0.

■

For $K = -\boldsymbol{\Delta}^{-1}$, the Osher-Solé-Vese problem is a particular case of the TV-Hilbert framework.

3.2 Mathematical study

Proposition 3.2 (Existence and Uniqueness)**.** *Let $\boldsymbol{f} \in L^2(\Omega,\mathbb{R}^M)$. The minimization problem :*

$$\inf\left\{|\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\langle \boldsymbol{f}-\boldsymbol{u}, K(\boldsymbol{f}-\boldsymbol{u})\rangle_{L^2(\Omega,\mathbb{R}^M)},\ \boldsymbol{u} \in BV(\Omega,\mathbb{R}^M),\ (\boldsymbol{f}-\boldsymbol{u}) \in V_0\right\}$$

has a unique solution $\boldsymbol{u} \in BV(\Omega,\mathbb{R}^M)$.

Proof:

- Existence :

  Let $E(\boldsymbol{u})$ denote the functional defined on $L^2(\Omega, \mathbb{R}^M)$ (with $E(u) = +\infty$ if $\boldsymbol{u} \notin BV(\Omega, \mathbb{R}^M)$ or $(\boldsymbol{f} - \boldsymbol{u}) \notin V_0$).

  Let us notice that $E \neq +\infty$ since $\boldsymbol{f}_\Omega = \frac{1}{|\Omega|}\int_\Omega \boldsymbol{f}$ belongs to $BV(\Omega, \mathbb{R}^M)$ and $(\boldsymbol{f} - \boldsymbol{f}_\Omega) \in V_0$.

  The functional E is convex. Since K is bounded, we deduce that E is lower semi-continuous for the $L^2(\Omega)$ weak topology. E is coercive : by the Poincaré inequality :

  $$\exists C > 0,\ \|\boldsymbol{u} - \boldsymbol{u}_\Omega\|_2 \leq C|\boldsymbol{u}|_{TV}$$

  with $\boldsymbol{u}_\Omega = \frac{1}{|\Omega|}\int_\Omega \boldsymbol{u} = \frac{1}{|\Omega|}\int_\Omega \boldsymbol{f}$ for $E(u) < +\infty$. Thus E has a minimizer.

- Uniqueness :

  We notice that the second term of the functional is strictly convex: it is the square norm in a Hilbert space.

  Since the first term is convex, the functional is strictly convex : the minimizer is unique.

■

We introduce the notation $\boldsymbol{v} = \boldsymbol{f} - \boldsymbol{u}$, when $\boldsymbol{u}$ is the unique minimizer of the TV-Hilbert problem.

Theorem 3.1 (Characterization of minimizers)**.** *Let $\boldsymbol{f} \in L^2(\Omega, \mathbb{R}^M)$.*

(i) If $\|K\boldsymbol{f}\|_* \leq \lambda$ then the solution of the TV-Hilbert problem is given by $(\boldsymbol{u}, \boldsymbol{v}) = (0, \boldsymbol{f})$.

(ii) If $\|K\boldsymbol{f}\|_* > \lambda$ then the solution $(\boldsymbol{u}, \boldsymbol{v})$ is characterized by:

$$\|K\boldsymbol{v}\|_* = \lambda \text{ and } \langle \boldsymbol{u}, K\boldsymbol{v}\rangle_{L^2(\Omega,\mathbb{R}^M)} = \lambda|\boldsymbol{u}|_{TV}.$$

Proof:

(i) $(0, \boldsymbol{f})$ is a minimizer iff

$$\begin{aligned} \forall \boldsymbol{h} \in BV(\Omega, \mathbb{R}^M), \forall \epsilon \in \mathbb{R}, |\epsilon\boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{f} - \epsilon\boldsymbol{h}\|^2_{\mathcal{H}} &\geq \frac{1}{2\lambda}\|\boldsymbol{f}\|^2_{\mathcal{H}}, \\ \text{which is equivalent to } |\epsilon||\boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\epsilon^2\|\boldsymbol{h}\|^2_{\mathcal{H}} - \frac{1}{\lambda}\epsilon\langle \boldsymbol{f}, \boldsymbol{h}\rangle_{\mathcal{H}} &\geq 0. \end{aligned}$$

We can divide by $|\epsilon| \to 0$, and depending on the sign of ϵ we get :

$$\pm\langle \boldsymbol{f}, \boldsymbol{h}\rangle_{\mathcal{H}} \leq \lambda|\boldsymbol{h}|_{TV}.$$

If $(0, \boldsymbol{f})$ is a minimizer, we can take the supremum for $\boldsymbol{h} \in BV(\Omega, \mathbb{R}^M)$. By definition of the *-norm, we have : $\|K\boldsymbol{f}\|_* \leq \lambda$.

Conversely, if $\boldsymbol{f} \in L^2(\Omega, \mathbb{R}^M)$ is such that $\|\boldsymbol{K}\boldsymbol{f}\|_* \leq \lambda$, the second inequality is true , thus $(0, \boldsymbol{f})$ is a minimizer.

(ii) As before, let us characterize the extremum: $(\boldsymbol{u}, \boldsymbol{v})$ is a minimizer iff

$$\begin{aligned} \forall \boldsymbol{h} \in BV(\Omega, \mathbb{R}^M), \forall \epsilon \in \mathbb{R}, |\boldsymbol{u} + \epsilon \boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{v} - \epsilon \boldsymbol{h}\|^2_{\mathcal{H}} &\geq |\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{v}\|^2_{\mathcal{H}}, \\ \text{or } |\boldsymbol{u} + \epsilon \boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\epsilon^2\|\boldsymbol{h}\|^2_{\mathcal{H}} - \frac{1}{\lambda}\epsilon\langle \boldsymbol{v}, \boldsymbol{h}\rangle_{\mathcal{H}} &\geq |\boldsymbol{u}|_{TV}. \end{aligned}$$

By the triangle inequality, we have:

$$\begin{aligned} |\boldsymbol{u}|_{TV} + |\epsilon||\boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\epsilon^2\|\boldsymbol{h}\|^2_{\mathcal{H}} - \frac{1}{\lambda}\epsilon\langle \boldsymbol{v}, \boldsymbol{h}\rangle_{\mathcal{H}} &\geq |\boldsymbol{u}|_{TV} \\ |\boldsymbol{h}|_{BV} &\geq \frac{1}{\lambda}\langle \boldsymbol{v}, \boldsymbol{h}\rangle_{\mathcal{H}} \\ \text{Taking the supremum } : \|K\boldsymbol{v}\|_* &\leq \lambda. \end{aligned}$$

Moreover, choosing $\boldsymbol{h} = \boldsymbol{u}$, $\epsilon \in]-1, 1[$:

$$\begin{aligned} (1+\epsilon)|\boldsymbol{u}|_{BV} &\geq \frac{1}{\lambda}\epsilon\langle \boldsymbol{v}, \boldsymbol{u}\rangle_{\mathcal{H}} + |\boldsymbol{u}|_{TV} \\ \text{For } \epsilon > 0 : \ |\boldsymbol{u}|_{TV} &\geq \frac{1}{\lambda}\langle \boldsymbol{v}, \boldsymbol{u}\rangle_{\mathcal{H}} \\ \text{For } \epsilon < 0 : \ |\boldsymbol{u}|_{TV} &\leq \frac{1}{\lambda}\langle \boldsymbol{v}, \boldsymbol{u}\rangle_{\mathcal{H}} \end{aligned}$$

We deduce that $\|K\boldsymbol{v}\|_*|\boldsymbol{u}|_{TV} \geq \langle \boldsymbol{v}, \boldsymbol{u}\rangle_{\mathcal{H}} = \lambda|\boldsymbol{u}|_{TV}$, and by the first upper-bound inequality, we have $\|K\boldsymbol{v}\|_* = \lambda$.

Conversely, let us assume these equalities hold. Then :

$$\begin{aligned} |\boldsymbol{u} + \epsilon \boldsymbol{h}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{v} - \epsilon \boldsymbol{h}\|^2_{\mathcal{H}} &\geq \frac{1}{\lambda}\langle (\boldsymbol{u} + \epsilon \boldsymbol{h}), K\boldsymbol{v}\rangle_{L^2(\Omega, \mathbb{R}^M)} + \frac{1}{2\lambda}\|\boldsymbol{v}\|^2_{\mathcal{H}} + \frac{1}{2\lambda}\|\boldsymbol{h}\|^2_{\mathcal{H}}\epsilon^2 - \frac{1}{\lambda}\epsilon\langle h, v\rangle_{\mathcal{H}} \\ &\geq |\boldsymbol{u}|_{TV} + \frac{1}{2\lambda}\|\boldsymbol{v}\|^2_{\mathcal{H}}. \end{aligned}$$

■

The mapping $v \mapsto \sup_{|\boldsymbol{u}|_{TV} \neq 0} \frac{\langle \boldsymbol{u}, K\boldsymbol{v}\rangle^2_{L^2(\Omega, \mathbb{R}^M)}}{|\boldsymbol{u}|_{TV}}$ is convex, lower semi-continuous for the $\mathcal{H}$ strong topology as a supremum of convex lower semi-continuous functions. As a consequence, for $\lambda > 0$ the set

$$G_\lambda = \{\boldsymbol{v} \in L^2(\Omega, \mathbb{R}^M), \|\boldsymbol{v}\|_* \leq \lambda\}$$

is a closed convex set, as well as $K^{-1}G_\lambda$. The orthogonal projection of this set is well-defined and we can notice that Theorem 3.1 reformulates :

$$\left\{ \begin{array}{rcl} \boldsymbol{v} &=& P^{\mathcal{H}}_{K^{-1}G_\lambda}(\boldsymbol{f}) \\ \boldsymbol{u} &=& \boldsymbol{f} - \boldsymbol{v}. \end{array} \right.$$

Indeed, if $(\boldsymbol{u}, \boldsymbol{v})$ is a minimizer of the TV-Hilbert problem, with $\boldsymbol{f} = \boldsymbol{u} + \boldsymbol{v}$, we have $\boldsymbol{v} \in K^{-1}G_\lambda$ and for any $\tilde{\boldsymbol{v}} \in K^{-1}G_\lambda$,

$$\langle \boldsymbol{f} - \boldsymbol{v}, \tilde{\boldsymbol{v}} - \boldsymbol{v}\rangle_{\mathcal{H}} = \langle \boldsymbol{u}, K\tilde{\boldsymbol{v}}\rangle_{L^2(\Omega, \mathbb{R}^M)} - \langle \boldsymbol{u}, K\boldsymbol{v}\rangle_{L^2(\Omega, \mathbb{R}^M)} \leq \|K\tilde{\boldsymbol{v}}\|_*|\boldsymbol{u}|_{TV} - \lambda|\boldsymbol{u}|_{TV} \leq 0,$$

thus by the equivalent definition of the projection on a closed convex set (see Ekeland-Témam [35]), we obtain the desired result.

Consequently, $\boldsymbol{v}$ is the orthogonal projection of $\boldsymbol{f}$ on the set $K^{-1}G_\lambda$ where

$$G_\lambda = \{\lambda \text{div } \vec{\boldsymbol{p}}, |\vec{\boldsymbol{p}}| \leq 1\}$$

(see Theorem 2.2), and the problem is equivalent to its dual formulation:

$$\inf_{|\vec{\boldsymbol{p}}|\leq 1} \|\lambda K^{-1} \text{ div } \vec{\boldsymbol{p}} - \boldsymbol{f}\|^2_{\mathcal{H}}. \tag{11}$$

4 Projected gradient Algorithm

Fast algorithms can be obtained by solving the dual formulation (11). For grayscale images, the famous projection algorithm by A. Chambolle [20] was the inspiration for all the following algorithms. We present here a projection algorithm, and we provide a complete proof of convergence of this scheme. We note that an independent work has just been reported in [43], where the authors M. Zhu, S.J. Wright, and T.F Chan have also applied the projected gradient method for solving the dual formulation of total variation minimization for image denoising They have a general framework also, although applied only to scalar image denoising and not related to image decompositions.

4.1 Discrete setting

From now on, we will work in the discrete case, using the following convention. A grayscale image is a matrix of size $N \times N$. We write $X = \mathbb{R}^{N\times N}$ the space of grayscale images. Their gradients belong to the space $Y = X \times X$. The L^2 inner product is $\langle u, v\rangle_X = \sum_{1\leq i,j\leq N} u_{i,j}v_{i,j}$.

The gradient operator is defined by $(\nabla u)_{i,j} = ((\nabla u)^x_{i,j}, (\nabla u)^y_{i,j})$ with:

$$(\nabla u)^x_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } i < N \\ 0 & \text{if } i = N \end{cases} \quad \text{and } (\nabla u)^y_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } j < N \\ 0 & \text{if } j = N. \end{cases}$$

The divergence operator is defined as the opposite of the adjoint operator of ∇:

$$\forall \vec{p} \in Y, \langle -\text{div } \vec{p}, u\rangle_X = \langle \vec{p}, \nabla u\rangle_Y$$

$$(\text{div } \vec{p})_{i,j} = \begin{cases} p^x_{i+1,j} - p^x_{i,j} & \text{if } 1 < i < N \\ p^x_{i,j} & \text{if } i = 1 \\ -p^x_{i-1,j} & \text{if } i = N \end{cases} + \begin{cases} p^x_{i,j+1} - p^y_{i,j} & \text{if } 1 < j < N \\ p^y_{i,j} & \text{if } j = 1 \\ -p^y_{i,j-1} & \text{if } j = N. \end{cases}$$

A color image is an element of X^M. The gradient and the divergence are defined component by component, and the L^2 inner product is given by :

$$\forall \boldsymbol{u}, \boldsymbol{v} \in X^M, \langle \boldsymbol{u}, \boldsymbol{v}\rangle_{X^M} = \sum_{k=1}^{M} \langle u^{(k)}, v^{(k)}\rangle_X$$

$$\forall \vec{\boldsymbol{p}}, \vec{\boldsymbol{q}} \in Y^M, \langle \vec{\boldsymbol{p}}, \vec{\boldsymbol{q}} \rangle_{Y^M} = \sum_{k=1}^{M} \langle p^{(k)}, q^{(k)} \rangle_Y$$

so that the color divergence is still the opposite of the adjoint of the color gradient.

4.2 Bresson-Chan algorithm

Problem (11) for grayscale images was tackled in [20] in the case $K = Id$, then in [13] for a general K. For color images, X. Bresson and T. Chan [19] showed that Chambolle's projection algorithm could still be used when $K = Id$. It is actually easy to check that it can be used with a general K for color images as well.

Following the steps of [13], one can notice that, provided $\tau \leq \frac{1}{8\|K^{-1}\|_{L^2}}$, the fixed point iteration:

$$\vec{\boldsymbol{p}}^{\,n+1} = \frac{\vec{\boldsymbol{p}}^{\,n} + \tau(\boldsymbol{\nabla}(K^{-1}\text{div } \vec{\boldsymbol{p}}^{\,n} - \boldsymbol{f}/\lambda)}{1 + \tau|\boldsymbol{\nabla}(K^{-1}\text{div } \vec{\boldsymbol{p}}^{\,n} - \boldsymbol{f}/\lambda)|} \tag{12}$$

gives a sequence $(\vec{\boldsymbol{p}}^{\,n})_{n\in\mathbb{N}}$ such that $\lambda K^{-1}\text{div } \vec{\boldsymbol{p}}^{\,n+1} \to v_\lambda$, and $f - \lambda K^{-1}\text{div } \vec{\boldsymbol{p}}^{\,n+1} \to u_\lambda$ when $n \to +\infty$.

Notice that the upper bound on τ is the same as for grayscale images.

4.3 Projected gradient

It was recently noticed ([21], [12]), that Problem (11) for grayscale images could be solved using a projected gradient descent. This is the algorithm we decided to extend to the case of color images.

Let $\mathcal{B}$ be the discrete version of our set of test-functions :

$$\mathcal{B} = \{\boldsymbol{v} \in Y^M, \forall\ 1 \leq i,j \leq N, |v_{i,j}|_2 \leq 1\}.$$

The orthogonal projection on $\mathcal{B}$ is easily computed:

$$P_{\mathcal{B}}(x) = \left(\frac{x_1}{\max\{1, |x|_2\}}, \frac{x_2}{\max\{1, |x|_2\}}\right).$$

The projected gradient descent scheme is defined by :

$$\vec{\boldsymbol{p}}^{m+1} = P_{\mathcal{B}}\left(\vec{\boldsymbol{p}}^{m+1} + \tau\nabla(K^{-1}\text{div } \vec{\boldsymbol{p}}^{m} - \boldsymbol{f}/\lambda\right) \tag{13}$$

which amounts to:

$$\boldsymbol{p}_{i,j}^{m+1} = \frac{\boldsymbol{p}_{i,j}^{m} + \tau\nabla(K^{-1}\text{div } \boldsymbol{p}^m - \frac{\boldsymbol{f}}{\lambda})_{i,j}}{\max\left(1, |\boldsymbol{p}_{i,j}^{m} + \tau\nabla(K^{-1}\text{div } \boldsymbol{p}^m - \frac{\boldsymbol{f}}{\lambda})_{i,j}|_2\right)} \tag{14}$$

where $\tau = \frac{\lambda}{\mu}$.

The convergence result for the projected gradient descent in the case of elliptic functions is standard (see [27], Theorem 8.6-2). Yet in our case the functional is not elliptic, and the proof needs a little more work.

Proposition 4.1. *If* $0 < \tau < \frac{1}{4\|K^{-1}\|}$*, then algorithm (14) converges. More precisely, there exists* $\vec{\tilde{\boldsymbol{p}}} \in \mathcal{B}$ *such that :*

$$\lim_{m\to\infty}(K^{-1}\ div\ \vec{\boldsymbol{p}}^m) = K^{-1}\ div\ \vec{\tilde{\boldsymbol{p}}}$$

and

$$\|\lambda K^{-1}\ div\ \vec{\tilde{\boldsymbol{p}}} - \boldsymbol{f}\|^2_{\mathcal{H}} = \inf_{\vec{\boldsymbol{p}}\in\mathcal{B}} \|\lambda K^{-1}\ div\ \vec{\boldsymbol{p}} - \boldsymbol{f}\|^2_{\mathcal{H}}$$

Proof: Writing

$\|K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{q}} - \boldsymbol{f}/\lambda\|^2_{\mathcal{H}} = \|K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}} - \boldsymbol{f}/\lambda\|^2_{\mathcal{H}} + \langle KK^{-1}\ \mathrm{div}\ (\vec{\boldsymbol{q}} - \vec{\boldsymbol{p}}), K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}} - \boldsymbol{f}/\lambda\rangle_{L^2} + O(\|\vec{\boldsymbol{q}} - \vec{\boldsymbol{p}}\|^2)$,

we begin by noticing that $\vec{\boldsymbol{p}}$ is a minimizer iff:

$$\vec{\boldsymbol{p}} \in \mathcal{B} \text{ and } \forall \vec{\boldsymbol{q}} \in \mathcal{B}, \forall \tau > 0, \langle \vec{\boldsymbol{q}} - \vec{\boldsymbol{p}}, \vec{\boldsymbol{p}} - (\vec{\boldsymbol{p}} + \tau\nabla(K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}} - \boldsymbol{f}/\lambda))\rangle_{L^2} \geq 0$$

Or equivalently:

$$\vec{\boldsymbol{p}} = P_{\mathcal{B}}\left(\vec{\boldsymbol{p}} + \tau(\nabla(K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}} - \boldsymbol{f}/\lambda)\right)$$

where $P_{\mathcal{B}}$ is the orthogonal projection on $\mathcal{B}$ with respect to the L^2 inner product.

We know that such a minimizer exists. Let us denote it by $\vec{\boldsymbol{p}}$.

- Now let us consider a sequence defined by (13), and write $A = -\nabla K^{-1}\ \mathrm{div}$. We have :

$$\begin{aligned}
\|\vec{\boldsymbol{p}}^{k+1} - \vec{\boldsymbol{p}}\|^2 &= \|P_{\mathcal{B}}(\vec{\boldsymbol{p}}^k + \tau\nabla(K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}}^k - \boldsymbol{f}/\lambda)) - P_{\mathcal{B}}(\vec{\boldsymbol{p}} + \tau\nabla(K^{-1}\ \mathrm{div}\ \vec{\boldsymbol{p}} - \boldsymbol{f}/\lambda))\|^2 \\
&\leq \|\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k + \tau\nabla K^{-1}\ \mathrm{div}\ (\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k)\|^2 \text{ since } P_{\mathcal{B}} \text{ is 1-Lipschitz [27]} \\
&\leq \|(I - \tau A)(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k)\|^2
\end{aligned}$$

Provided $\|I - \tau A\| \leq 1$, we can deduce :

$$\|\vec{\boldsymbol{p}}^{k+1} - \vec{\boldsymbol{p}}\| \leq \|\vec{\boldsymbol{p}}^k - \vec{\boldsymbol{p}}\| \tag{15}$$

and the sequence $(\|\vec{\boldsymbol{p}}^k - \vec{\boldsymbol{p}}\|)$ is convergent.

- A is a symmetric positive semi-definite operator. By writing $E = \ker A$ and $F = ImA$, we have:

$$Y^M = E \overset{\perp}{\oplus} F$$

and we can decompose any $\vec{\boldsymbol{q}} \in Y^M$ as the sum of two orthogonal components $\vec{\boldsymbol{q}}_E \in E$ and $\vec{\boldsymbol{q}}_F \in F$. Notice that by injectivity of K^{-1}, E is actually equal to the kernel of the divergence operator.

Let $\mu_1 = 0 < \mu_2 \leq \ldots \leq \mu_a$ be the ordered eigenvalues of A.

$$\begin{aligned}\|I-\tau A\| &= \max(|1-\tau\mu_1|,|1-\tau\mu_a|)\\ &= \max(1,|1-\tau\mu_a|)\\ &= 1 \text{ for } 0\leq\tau\leq\frac{2}{\mu_a}\end{aligned}$$

We can restrict $I-\tau A$ to F and then define :

$$\begin{aligned}g(\tau)=\|(I-\tau A)_{|F}\| &= \max(|1-\tau\mu_2|,|1-\tau\mu_a|)\\ &< 1 \quad \text{for } 0<\tau<\frac{2}{\mu_a}\end{aligned}$$

- Now we assume that $0<\tau<\frac{2}{\mu_a}$. Therefore, inequality (15) is true and the sequence $(\vec{\boldsymbol{p}}^k)$ is bounded, and so is the sequence $(K^{-1}\text{ div }\vec{\boldsymbol{p}}^k)$. We are going to prove that the sequence $(K^{-1}\text{ div }\vec{\boldsymbol{p}}^k)$ has a unique cluster point. Let $(K^{-1}\text{ div }\vec{\boldsymbol{p}}^{\varphi(k)})$ be a convergent subsequence. By extraction, one can assume that $\vec{\boldsymbol{p}}^{\varphi(k)}$ is convergent too, and denote by $\tilde{\vec{\boldsymbol{p}}}$ its limit.

  Passing to the limit in (13), the sequence $(\vec{\boldsymbol{p}}^{\varphi(k)+1})$ is convergent towards:

$$\hat{\vec{\boldsymbol{p}}}=P_{\mathcal{B}}\left(\tilde{\vec{\boldsymbol{p}}}+\tau\nabla(K^{-1}\text{div }\tilde{\vec{\boldsymbol{p}}}-\boldsymbol{f}/\lambda)\right)$$

  Using (15), we also notice that :

$$\|\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}}\|=\|\hat{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}}\|$$

  As a consequence:

$$\begin{aligned}\|\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}}\|^2 &= \|\hat{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}}\|^2\\ &= \|P_{\mathcal{B}}\left(\tilde{\vec{\boldsymbol{p}}}+\tau\nabla(K^{-1}\text{div }\tilde{\vec{\boldsymbol{p}}}-\boldsymbol{f}/\lambda)\right)-P_{\mathcal{B}}\left(\vec{\boldsymbol{p}}+\tau\nabla(K^{-1}\text{div }\vec{\boldsymbol{p}}-\boldsymbol{f}/\lambda)\right)\|^2\\ &\leq \|(I-\tau A)(\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})\|^2\\ &\leq \|(\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})_E\|^2+g(\tau)^2\|(\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})_F\|^2\\ &< \|\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}}\|^2 \text{ if } (\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})_F\neq 0\end{aligned}$$

  Of course, this last inequality cannot hold, which means that $\|(\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})_F\|=0$. Hence

$$(\tilde{\vec{\boldsymbol{p}}}-\vec{\boldsymbol{p}})\in E=\ker A \text{ and } K^{-1}\text{ div }\tilde{\vec{\boldsymbol{p}}}=K^{-1}\text{ div }\vec{\boldsymbol{p}}$$

  so the sequence $(K^{-1}\text{ div }\vec{\boldsymbol{p}}^k)$ is convergent.

- The last remark consists in evaluating μ_a. We have:

$$\mu_a=\|\nabla K^{-1}\text{ div }\|\leq\|\nabla\|\|K^{-1}\|\|\text{ div }\|$$

  Since $\|\text{ div }\|^2=\|\nabla\|^2=8$ (see [20], the result is still true for color images), we deduce that

$$\mu_a\leq 8\|K^{-1}\|$$

■

Since we are only interested in $\boldsymbol{v} = \lambda K^{-1}\mathrm{div}\,\vec{\boldsymbol{p}}$, Proposition (4.1) justifies the validity of algorithm (13). Yet it is possible to prove that $\vec{\boldsymbol{p}}$ itself converges.

Lemma 4.1. *Let P orthogonal projection on a nonempty closed convex set K. Let $Q = Id - P$. Then we have:*

$$\|Q(v_1) - Q(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \leq \|v_1 - v_2\|^2 \tag{16}$$

Proof:

$$\begin{aligned}
& \|Q(v_1) - Q(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \\
= \quad & \|v_1 - v_2 + P(v_1) - P(v_2)\|^2 + \|P(v_1) - P(v_2)\|^2 \\
= \quad & \|v_1 - v_2\|^2 + 2\|P(v_1) - P(v_2)\|^2 - 2\langle P(v_1) - P(v_2), v_1 - v_2\rangle \\
= \quad & \|v_1 - v_2\|^2 + 2\langle P(v_1) - P(v_2), P(v_1) - P(v_2) - v_1 + v_2\rangle \\
= \quad & \|v_1 - v_2\|^2 + 2\underbrace{\langle P(v_1) - P(v_2), P(v_1) - v_1\rangle}_{\leq 0} + 2\underbrace{\langle P(v_2) - P(v_1), P(v_2) - v_2\rangle}_{\leq 0}
\end{aligned}$$

(using the characterization of the projection on a nonempty closed convex set [27]).

■

Remark: We have:

$$\vec{\boldsymbol{p}} = P_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) = \vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) - Q_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda))$$

And thus:

$$-\tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) = Q_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) \tag{17}$$

Corollary 4.1. *The sequence $\vec{\boldsymbol{p}}^k$ defined by (13) converges to $\vec{\boldsymbol{p}}$ unique solution of problem (11).*

Proof: Using the above lemma, we have:

$$\begin{aligned}
& \|Q_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) - Q_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda))\|^2 \\
& \quad + \|P_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) - P_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda))\|^2 \\
\leq \quad & \|\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda) - \vec{\boldsymbol{p}}^k + \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda)\|^2 \\
\leq \quad & \|\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k\|^2 + \tau^2\|A(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k))\|^2 - 2\tau\langle \vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k, A(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k)\rangle
\end{aligned}$$

Hence:

$$\begin{aligned}
\|Q_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)) \quad - \quad & Q_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda))\|^2 + \|\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^{k+1}\|^2 \\
& \leq \|\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k\|^2 + \tau^2\|A(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k))\|^2 - 2\tau\langle \vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k, A(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k)\rangle
\end{aligned}$$

But we have already shown that $\|\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k\|$ converges, and that $K^{-1}\text{div}\,(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k) \to 0$ (therefore $A(\vec{\boldsymbol{p}} - \vec{\boldsymbol{p}}^k) \to 0$). Hence, passing to the limit in the above equation, we get:

$$Q(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda)) \to Q_{\mathcal{B}}(\vec{\boldsymbol{p}} - \tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda))$$

We thus deduce from (17) that

$$Q_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda)) \to -\tau(A\vec{\boldsymbol{p}} + \nabla \boldsymbol{f}/\lambda)$$

Remembering that $P_{\mathcal{B}} + Q_{\mathcal{B}} = Id$, we get:

$$\vec{\boldsymbol{p}}^{k+1} = P_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda)) = (Id - Q_{\mathcal{B}})(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda))$$

Hence:

$$\vec{\boldsymbol{p}}^{k+1} - \vec{\boldsymbol{p}}^k = -\tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda)) - Q_{\mathcal{B}}(\vec{\boldsymbol{p}}^k - \tau(A\vec{\boldsymbol{p}}^k + \nabla \boldsymbol{f}/\lambda))$$

Passing to the limit, we get:

$$\vec{\boldsymbol{p}}^{k+1} - \vec{\boldsymbol{p}}^k \to 0$$

We can now pass to the limit in (13) and get that $\vec{\boldsymbol{p}}^k \to \vec{\boldsymbol{p}}$.

■

5 Applications to color image denoising and decomposition

In this last section, we apply the projected gradient algorithm to solve various color image problems.

5.1 TV-Hilbert model

5.1.1 The color ROF model

As an application of (14), we use the following scheme for the ROF model (9):

$$\boldsymbol{p}_{i,j}^{m+1} = \frac{\boldsymbol{p}_{i,j}^m + \tau\nabla(\text{div}\ \boldsymbol{p}^m - \frac{\boldsymbol{f}}{\lambda})_{i,j}}{\max\left(1, |\boldsymbol{p}_{i,j}^m + \tau\nabla(\text{div}\ \boldsymbol{p}^m - \frac{\boldsymbol{f}}{\lambda})_{i,j}|_2\right)} \tag{18}$$

5.1.2 The color OSV model

As for the OSV model (10), we use:

$$\boldsymbol{p}_{i,j}^{m+1} = \frac{\boldsymbol{p}_{i,j}^m - \tau\nabla(\Delta\text{div}\ \boldsymbol{p}^m + \frac{\boldsymbol{f}}{\lambda})_{i,j}}{\max\left(1, |\boldsymbol{p}_{i,j}^m - \tau\nabla(\Delta\text{div}\ \boldsymbol{p}^m + \frac{\boldsymbol{f}}{\lambda})_{i,j}|_2\right)} \tag{19}$$

5.2 The color A2BC algorithm

Following Y. Meyer [44], one can use the $G(\Omega)$ space to model textures, and try to solve this problem:

$$\inf_{\boldsymbol{u}} |\boldsymbol{u}|_{BV} + \alpha \|\boldsymbol{f} - \boldsymbol{u}\|_{G} \tag{20}$$

In [7], J.-F. Aujol and G. Aubert , L. Blanc-Féraud and A. Chambolle approximate this problem by minimizing the following functional:

$$F_{\mu,\lambda}(\boldsymbol{u},\boldsymbol{v}) = \begin{cases} |\boldsymbol{u}|_{BV} + \frac{1}{2\lambda}\|\boldsymbol{f} - \boldsymbol{u} - \boldsymbol{v}\|^2_{L^2(\Omega)} \text{ if } (\boldsymbol{u},\boldsymbol{v}) \in BV(\Omega,\mathbb{R}) \times G_\mu \\ +\infty \text{ otherwise} \end{cases} \tag{21}$$

or equivalently :

$$F_{\mu,\lambda}(\boldsymbol{u},\boldsymbol{v}) = |\boldsymbol{u}|_{BV} + \frac{1}{2\lambda}\|\boldsymbol{f} - \boldsymbol{u} - \boldsymbol{v}\|^2_{L^2(\Omega)} + \chi_{G_\mu}(\boldsymbol{v}_n)$$

with $\chi_{G_\mu}(\boldsymbol{v}) = \begin{cases} 0 \text{ if } \boldsymbol{v} \in G_\mu \\ +\infty \text{ otherwise} \end{cases}$

The generalization to color images was done by J.-F. Aujol and S. H. Kang in [10] using a chromaticity-brightness model. The authors used a gradient descent in order to compute the projections, which is rather slow, and requires to regularize the total variation.

In [19], X. Bresson and T. Chan used the following scheme (but relying on Chambolle's algorithm) for color images in order to compute the projections. As in the grayscale case, the minimization is done using an alternate scheme (but in the present paper we use the projected gradient descent scheme described before to compute the projections):

- Initialization:

$$\boldsymbol{u}_0 = \boldsymbol{v}_0 = 0$$

- Iterations:

$$\boldsymbol{v}_{n+1} = P_{G_\mu}(\boldsymbol{f} - \boldsymbol{u}_n) \tag{22}$$
$$\boldsymbol{u}_{n+1} = \boldsymbol{f} - \boldsymbol{v}_{n+1} - P_{G_\lambda}(\boldsymbol{f} - \boldsymbol{v}_{n+1}) \tag{23}$$

- Stop if the following condition is true:

$$\max(|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n|, |\boldsymbol{v}_{n+1} - \boldsymbol{v}_n|) \leq \epsilon$$

In [11], it is shown that under reasonable assumptions, the solutions of Problem (21) converge when $\lambda \to 0$ to a solution of Problem (20) for a certain α.

5.3 The color TV-L1 model

The TV-L1 model is very popular for grayscale images. It benefits from having both good theoretical properties (it is a morphological filter) and fast algorithms (see [30]). In order to extend it to color images, we consider the following problem:

$$\inf_{\boldsymbol{u}} |\boldsymbol{u}|_{TV} + \lambda \|\boldsymbol{f} - \boldsymbol{u}\|_1 \tag{24}$$

with the notation:

$$\|\boldsymbol{u}\|_1 = \int_\Omega \sqrt{\sum_{l=1}^{M} |u_l|^2}$$

(like for the total variation we have decided to have a coupling between channels).

Our approach is different from the one used by J. Darbon in [30], since it was using a channel by channel decomposition with the additional constraint that no new color is created. As for the A2BC algorithm, we are led to consider the approximation:

$$\inf_{\boldsymbol{u},\boldsymbol{v}} |\boldsymbol{u}|_{BV} + \frac{1}{2\alpha}\|\boldsymbol{f} - \boldsymbol{u} - \boldsymbol{v}\|_2^2 + \lambda\|\boldsymbol{v}\|_1$$

Once again, having a projection algorithm for color images allows us to generalize easily this problem. In order to generalize the TV-L1 algorithm proposed by J.-F. Aujol, G. Gilboa, T. Chan and S. Osher ([14]), we aim at solving the alternate minimization problem:

(i)

$$\inf_{\boldsymbol{u}} |\boldsymbol{u}|_{BV} + \frac{1}{2\alpha}\|\boldsymbol{f} - \boldsymbol{u} - \boldsymbol{v}\|_2^2$$

(ii)

$$\inf_{\boldsymbol{v}} \frac{1}{2\alpha}\|\boldsymbol{f} - \boldsymbol{u} - \boldsymbol{v}\|_2^2 + \lambda\|\boldsymbol{v}\|_1$$

The first problem is a Rudin-Osher-Fatemi problem. Scheme (14) with $K = Id$ is well adapted for solving it. For the second one, the following property shows that a "vectorial soft thresholding" gives the solution:

Proposition 5.1. *The solution of problem (ii), is given by:*

$$\boldsymbol{v}(x) = VT_{\alpha\lambda}(\boldsymbol{f}(x) - \boldsymbol{u}(x)) = \frac{\boldsymbol{f}(x) - \boldsymbol{u}(x)}{|\boldsymbol{f}(x) - \boldsymbol{u}(x)|_2} \max\left(|\boldsymbol{f}(x) - \boldsymbol{u}(x)|_2 - \alpha\lambda, 0\right) \textit{ almost everywhere}$$

The proof of this last result is a straightforward extension of Proposition 4 in [14].

Henceforth, we propose the following generalization of the TV-L1 algorithm:

- Initialization:

$$\boldsymbol{u}_0 = \boldsymbol{v}_0 = 0$$

- Iterations:

$$\begin{aligned} \boldsymbol{v}_{n+1} &= VT_{\alpha\lambda}(\boldsymbol{f} - \boldsymbol{u}_n) && (25)\\ \boldsymbol{u}_{n+1} &= \boldsymbol{f} - \boldsymbol{v}_{n+1} - P_{G_\alpha}(\boldsymbol{f} - \boldsymbol{v}_{n+1}) && (26) \end{aligned}$$

- Stop if the following condition is satisfied:

$$\max(|\boldsymbol{u}_{n+1} - \boldsymbol{u}_n|, |\boldsymbol{v}_{n+1} - \boldsymbol{v}_n|) \leq \epsilon$$

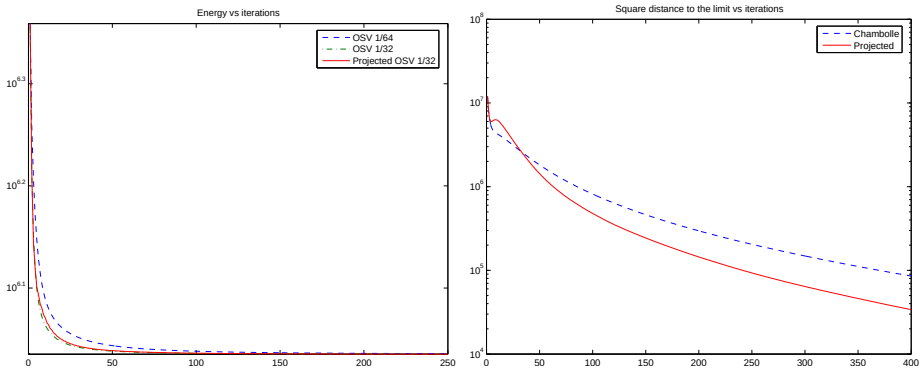

Figure 1: Left: Energy vs iterations of the Osher-Solé-Vese model with Chambolle's projection algorithm (in green and blue - stepsize 1/64 and 0.031) and with the Projected gradient algorithm (in red - stepsize 0.031). Right: L^2 square distance (on a logarithmic scale) between the limit value (2000 iterations) vs the number of iterations, for OSV using 1/64 stepsize.

5.4 Numerical experiments

Figure 1 shows the decrease of the energy and the convergence of the projected gradient algorithm for the OSV model (10). We compare scheme (14) with the projection algorithm of [8] (which is a straightforward modification of Chambolle's projection algorithm [20]). Both algorithms behave similarly and it is hard to tell whether one converges faster than the other. Figures 2 and 3 display denoising results using ROF (9) and OSV (10) models. The images look very similar but since the OSV model penalizes much more the highest frequencies than the ROF model [14], the denoised image still shows the lowest frequencies of the noise.

Figure 4 shows the noisy part using these experiments. As one could expect it is much more oscillating in the OSV model.

Figures 5 and 6 display a cartoon-texture decomposition experiment using different kinds of texture. The algorithm used were A2BC and TVL1. Both results look good.

On Figure 7, a denoising experiment was performed using salt-and-pepper noise. The denoised picture looks quite good and surprisingly better than the original image! This is because the picture we used had some compression artifacts that the algorithm removed.

References

[1] R. Acar and C. Vogel. Analysis of total variation penalty methods for ill-posed problems. *Inverse Problems*, 10:1217–1229, 1994.

[2] R.A. Adams. *Sobolev Spaces*, volume 65 of *Mathematics*. Academic Press, 1975.

[3] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford mathematical monographs. Oxford University Press, 2000.

[4] F. Andreu-Vaillo, V. Caselles, and J. M. Mazon. *Parabolic quasilinear equations minimizing linear growth functionals*, volume 223 of *Progress in Mathematics*. Birkhauser, 2002.

Figure 2: First row: original and noisy images (PSNR = 56.9 dB). Second row: denoised with color ROF ($\lambda = 15$, PSNR= 73.8 dB) and with color OSV ($\lambda = 15$, PSNR= 73.4 dB).

[5] G. Aubert and J.F. Aujol. Modeling very oscillating signals. application to image processing. *Applied Mathematics and Optimization*, 51(2):163–182, mars 2005.

[6] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, volume 147. Springer Verlag, Applied Mathematical Sciences, 2001.

[7] J-F. Aujol, G. Aubert, L. Blanc-Féraud, and A. Chambolle. Image decomposition into a bounded variation component and an oscillating component. *Journal of Mathematical Imaging and Vision*, 22(1):71–88, January 2005.

[8] J-F. Aujol and A. Chambolle. Dual norms and image decomposition models. *International Journal on Computer Vision*, 63(1):85–104, June 2005.

[9] J-F. Aujol and T. F. Chan. Combining geometrical and textured information to perform image classification. *Journal of Visual Communication and Image representation*, 17(5):1004–1023, October 2006.

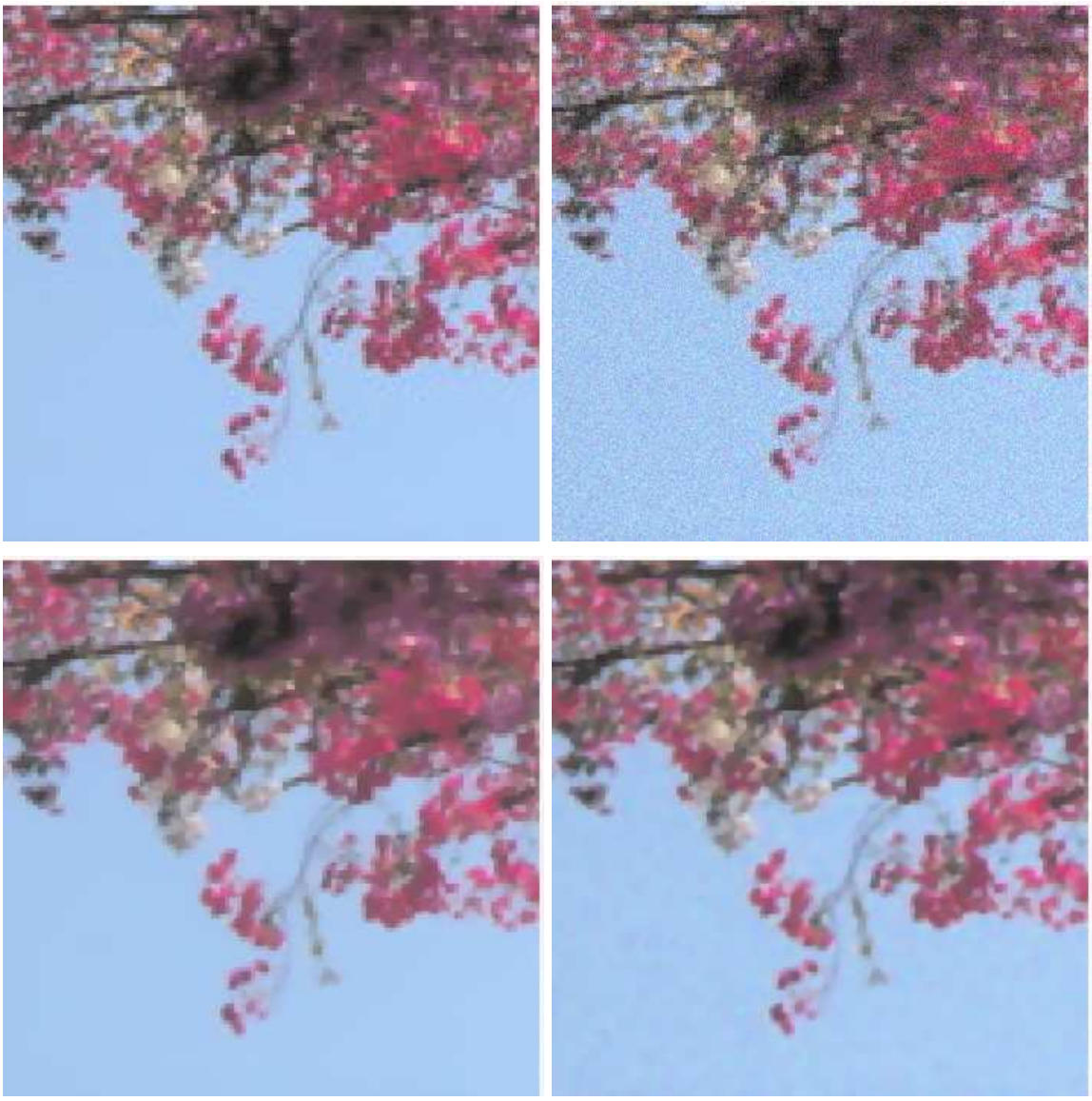

Figure 3: First row: original and noisy images (PSNR = 57.3 dB). Second row: denoised with color ROF ($\lambda = 25$, PSNR= 74.2 dB) and with color OSV ($\lambda = 25$, PSNR= 74.1 dB).

[10] J.-F. Aujol and S. H. Kang. Color image decomposition and restoration. *Journal of Visual Communication and Image Representation*, 17(4):916–928, August 2006.

[11] J.F. Aujol. *Contribution à l'Analyse de Textures en Traitement d'Images par Méthodes Variationnelles et Equations aux Dérivées Partielles*. PhD thesis, Universite de Nice Sophia Antipolis, juin 2004.

[12] J.F. Aujol. Some algorithms for total variation based image restoration. *CMLA Preprint 2008-05*, 2008. http://hal.archives-ouvertes.fr/hal-00260494/en/.

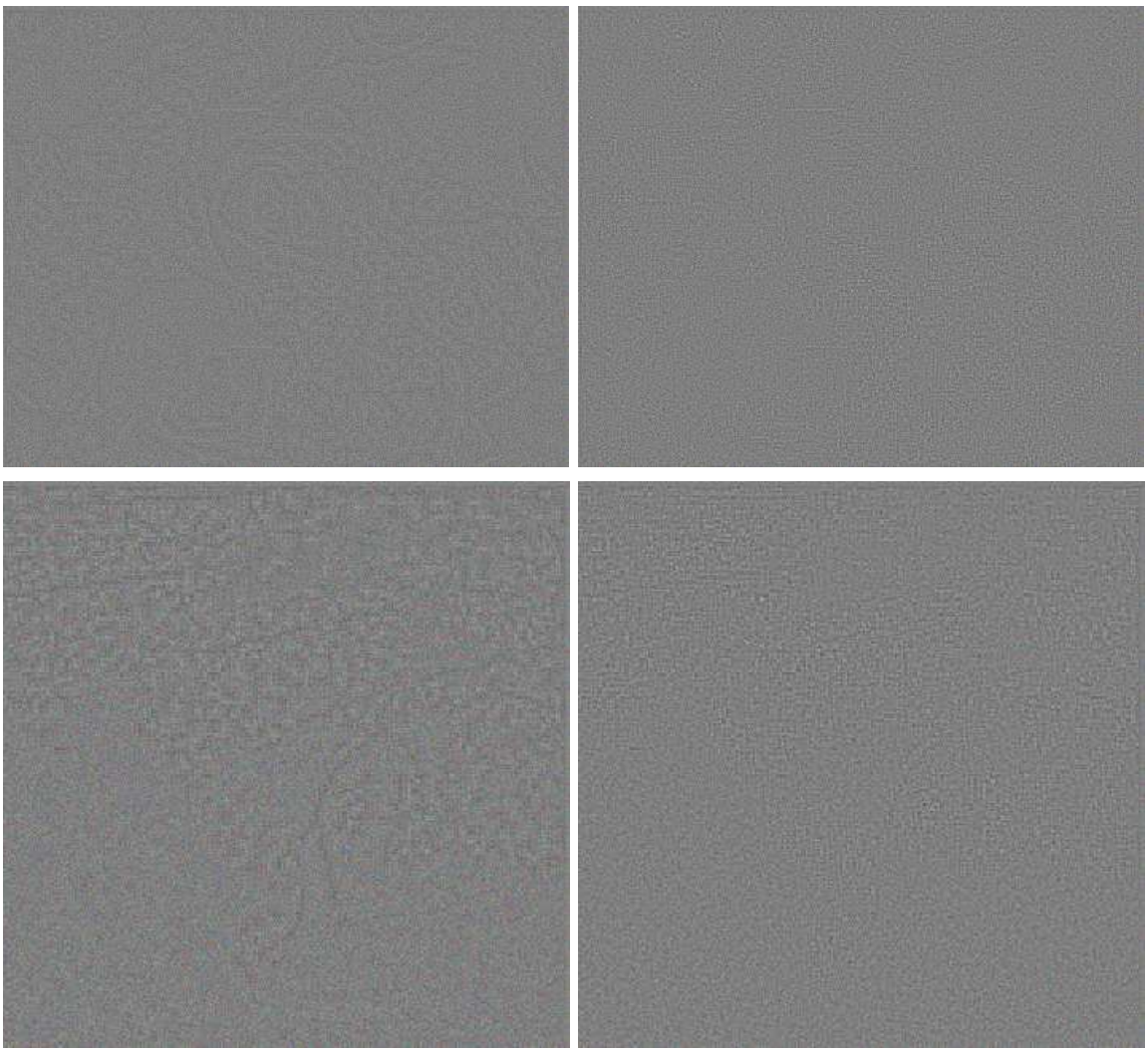

Figure 4: Noisy part (left column : color ROF, right column: color OSV).

[13] J.F. Aujol and G. Gilboa. Constrained and SNR-based solutions for TV-Hilbert space image denoising. *Journal of Mathematical Imaging and Vision*, 26(1-2):217–237, 2006.

[14] J.F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-texture image decomposition - modeling, algorithms, and parameter selection. *International Journal of Computer Vision*, 67(1):111–136, April 2006.

[15] J. Bect, L. Blanc-Féraud, G. Aubert, and A. Chambolle. A l1-unified variational framework for image restoration. In *ECCV 04*, volume 3024 of *Lecture Notes in Computer Sciences*, pages 1–13, 2004.

[16] A. Bermudez and C. Moreno. Duality methods for solving variational inequalities. *Comp. and Maths. with Appls.*, 7(1):43–58, July 1981.

Figure 5: Original images

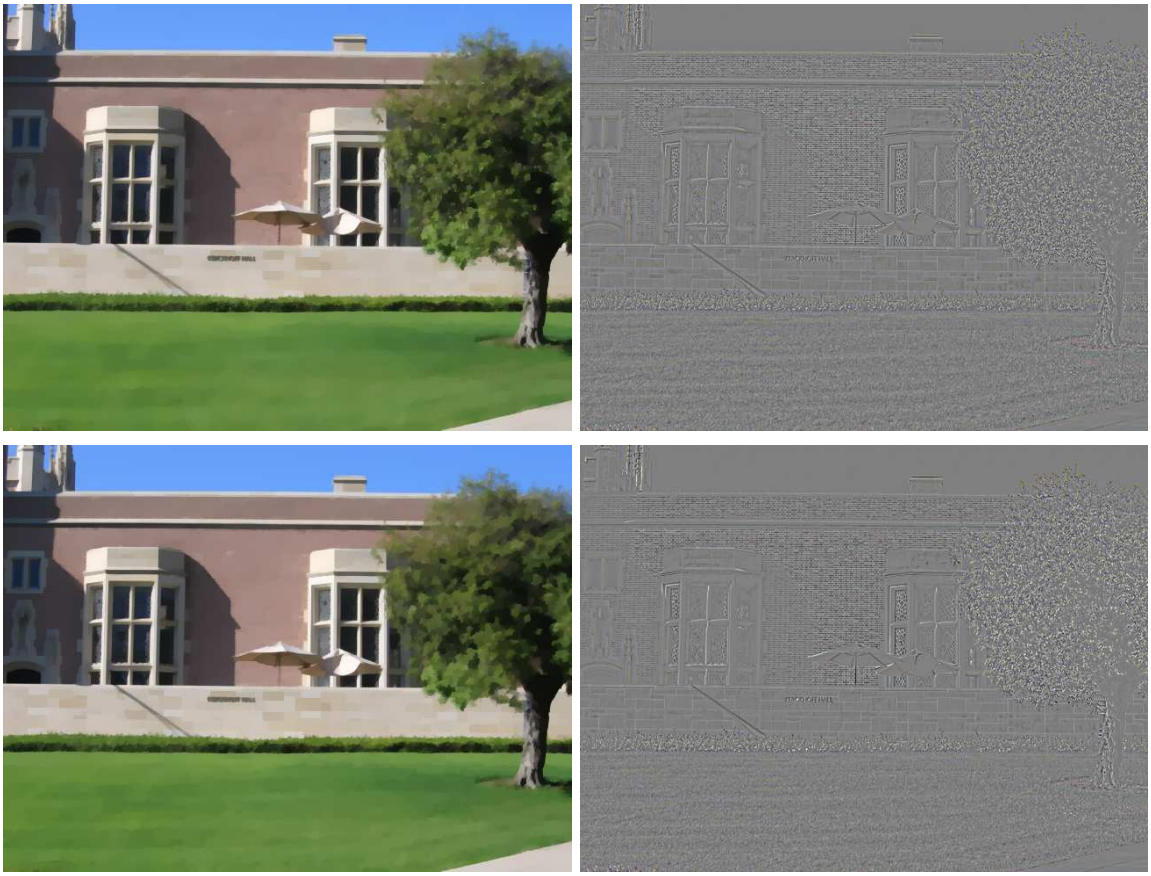

Figure 6: Cartoon-texture decomposition using color A2BC algorithm (upper row) and color TVL1 (lower row).

[17] J. Bioucas-Dias and M. Figueiredo. Thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, 16(12):2980–2991, 2007.

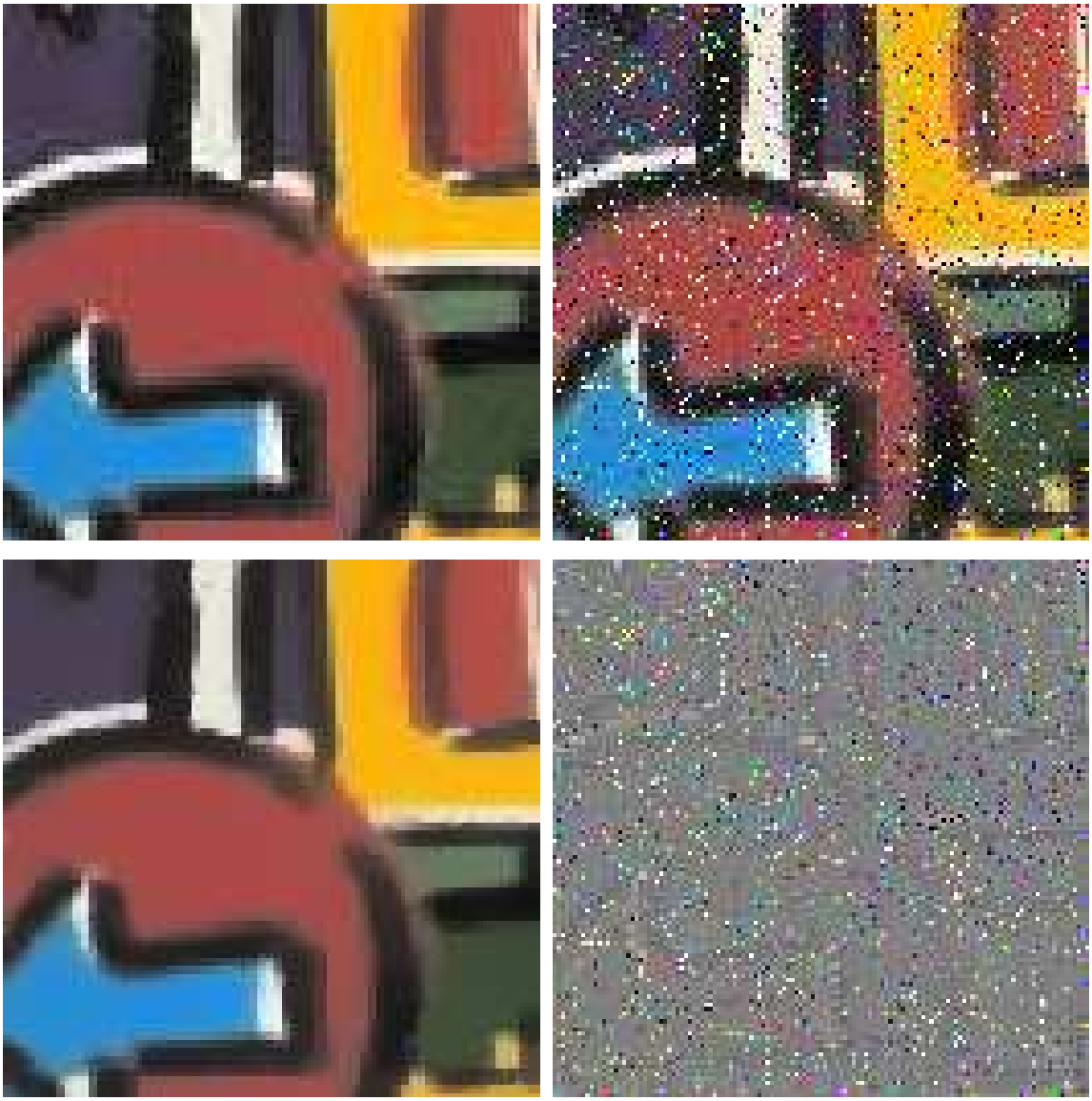

Figure 7: First row: original and noisy images (using salt and pepper noise, PSNR= 34.6 dB). Second row: denoised with color TVL1 (PSNR= 67.5 dB) and noise part.

[18] P. Blomgren and T. Chan. Color TV: Total variation methods for restoration of vector valued images. *IEEE Transactions on Image Processing*, 7(3):304–309, 1998.

[19] X. Bresson and T. Chan. Fast minimization of the vectorial total variation norm and applications to color image processing. *SIAM Journal on Imaging Sciences (SIIMS)*, Submitted, 2007.

[20] A. Chambolle. An algorithm for total variation minimization and its applications. *JMIV*, 20:89–97, 2004.

[21] A. Chambolle. Total variation minimization and a class of binary MRF models. *EMMCVPR in Vol. 3757 of Lecture Notes in Computer Sciences*, pages 136–152, June 2005.

[22] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(3):167–188, 1997.

[23] T. Chan and S. Esedoglu. Aspects of total variation regularized L^1 function approximation. *SIAM Journal of Applied Mathematics*, 65(5):1817–1837, 2005.

[24] T. Chan, G. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20(6):1964–1977, 1999.

[25] T. Chan and J. Shen. *Image processing and analysis - Variational, PDE, wavelet, and stochastic methods*. SIAM Publisher, 2005.

[26] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2), 2007.

[27] P. G Ciarlet. *Introduction à l'Analyse Numérique Matricielle et à l'Optimisation*. Dunod, 1998.

[28] P.L. Combettes and J. Pesquet. Image restoration subject to a total variation constraint. *IEEE Transactions on Image Processing*, 13(9):1213–1222, 2004.

[29] P.L. Combettes and V. Wajs. Signal recovery by proximal forward-backward splitting. *SIAM Journal on Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005.

[30] J. Darbon. Total variation minimization with l^1 data fidelity as a contrast invariant filter. *4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, pages 221–226, September 2005.

[31] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation part I: Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3):277–291, 2006.

[32] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*, 57:1413–1457, 2004.

[33] I. Daubechies and G. Teschke. Variational image restoration by means of wavelets: simultaneous decomposition, deblurring and denoising. *Applied and Computational Harmonic Analysis*, 19:1–16, 2005.

[34] D. Dobson and C. Vogel. Convergence of an iterative method for total variation denoising. *SIAM Journal on Numerical Analysis*, 34:1779–1791, 1997.

[35] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*, volume 28 of *Classics in Applied Mathematics*. SIAM, 1999.

[36] H. Fu, M. Ng, M. Nikolova, and J. Barlow. Efficient minimization methods of mixed l1-l1 and l2-l1 norms for image restoration. *SIAM Journal on Scientific computing*, 27(6):1881–1902, 2006.

[37] J. Garnett, P. Jones, T. Le, and L. Vese. Modeling oscillatory components with the homogeneous spaces $BMO^{-\alpha}$ and $W^{-\alpha,p}$. *UCLA CAM Report*, 07-21, July 2007.

[38] J. Gilles. Noisy image decomposition: a new structure, textures and noise model based on local adaptivity. *JMIV*, 28(3), July 2007.

[39] D. Goldfarb and W. Yin. Second-order cone programming methods for total variation based image restoration. *SIAM Journal on Scientific Computing*, 27(2):622–645, 2005.

[40] A. Haddad. *Méthodes variationnelles en traitement d'images*. PhD thesis, ENS Cachan, June 2005.

[41] T. Le and L. Vese. Image decomposition using total variation and div(BMO). *Multiscale Modeling and Simulation, SIAM Interdisciplinary Journal*, 4(2):390–423, June 2005.

[42] L. Lieu. *Contribution to Problems in Image Restoration, Decomposition, and Segmentation by Variational Methods and Partial Differential Equations,*. PhD thesis, UCLA, June 2006.

[43] S.J. Wright M. Zhu and T.F. Chan. Duality-based algorithms for total variation image restoration. Technical report, UCLA CAM Report 08-33, 2008.

[44] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations*, volume 22 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2001. The fifteenth Dean Jacqueline B. Lewis memorial lectures.

[45] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming (A)*, 103(1):127–152, 2005.

[46] M.K. Ng, L. Qi, Y.F. Yang, and Y. Huang. On semismooth Newton methods for total variation minimization. *Journal of Mathematical Imaging and Vision*, 27:265–276, 2007.

[47] M. Nikolova and R. Chan. The equivalence of half-quadratic minimization and the gradient linearization iteration. *IEEE Transactions on Image Processing*, 16(6):1623–1627, 2007.

[48] S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision, and Graphic*. Springer, 2003.

[49] S. Osher, A. Sole, and L. Vese. Image decomposition and restoration using total variation minimization and the H^{-1} norm. *SIAM journal on Multiscale Modeling and Simulation*, 1(3):349–370, 2003.

[50] L. Rudin, S. Osher, and E. Fatemi. Non linear total variation based noise removal algorithms. *Physica D*, 60:259–268, 2003.

[51] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing*, 5(11):1582–1586, 1996.

[52] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7(3):310–318, 1998.

[53] J.L. Starck, M. ELad, and D.L. Donoho. Image decomposition via the combination of sparse representations and a variational approach, 2005.

[54] L.A. Vese and S. J. Osher. Modeling textures with total variation minimization and oscillating patterns in image processing. *Journal of Scientific Computing*, 19(1-3):553–572, December 2003.

[55] L.A. Vese and S.J. Osher. Color texture modeling and color image decomposition in a variational-pde approach ,. In *Proceedings of the Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC '06)*, pages 103 – 110. IEEE, 2006.

[56] C.R. Vogel. *Computational Methods for Inverse Problems*, volume 23 of *Frontiers in Applied Mathematics*. SIAM, 2002.

[57] P. Weiss, G. Aubert, and L. Blanc-Feraud. Efficient schemes for total variation minimization under constraints in image processing. *INRIA Research Report*, 6260, 2007. http://hal.inria.fr/docs/00/16/62/30/PDF/RR-6260.pdf.

[58] W. Yin and D. Goldfarb. Second-order cone programming methods for total variation based image restoration. *SIAM J. Scientific Computing*, 27(2):622–645, 2005.

[59] W. Yin, D. Goldfarb, and S. Osher. Image cartoon-texture decomposition and feature selection using the total variation regularized l^1. In *Variational, Geometric, and Level Set Methods in Computer Vision*, volume 3752 of Lecture Notes in Computer Science, pages 73–84. Springer, 2005.

[60] W. Yin, D. Goldfarb, and S. Osher. A comparison of three total variation based texture extraction models. *Journal of Visual Communication and Image Representation*, 18(3):240–252, June 2007.